Inclusive Refugee-Hosting Improves Local Development and Prevents Public Backlash

December 16, 2022

For Online Publication: Supplementary Information

Contents

S1 Detailed Data Description **2**
S1.1 Unit of Analysis: Parish-Data Construction . . . 2
S1.2 Refugee Presence . . . 7
S1.3 Primary Schools . . . 11
S1.4 Secondary Schools . . . 12
S1.5 Health Access and Health Utilization (DHS) . . . 14
S1.6 Road Density . . . 17
S1.7 Afrobarometer Survey Data . . . 19
S1.8 ACLED Violence data . . . 20

S2 Robustness Checks **21**
S2.1 Alternative Refugee Presence Measures . . . 21
S2.2 Alternative Sample Radii . . . 24
S2.3 Including a District Time Trend . . . 28
S2.4 Two Period Binary Treatment DiD . . . 28

S3 Sensitivity Analysis **29**

S4 Multiple Hypothesis Testing Corrections **30**
S4.1 Public Goods Outcomes . . . 30
S4.2 Migration Attitudes (Afrobarometer) . . . 30
S4.3 Perceptions of Insecurity (Afrobarometer) . . . 30

S5 Addressing Compositional Concerns of Domestic Migration **31**

S1 Detailed Data Description

In this section, we provide detailed information on our procedure for matching administrative units across years, and for constructing and validating key variables used in the empirical analyses reported in the main paper. We begin with an exhaustive discussion of how parishes are matched and merged successfully across years in the face of a massive administrative change that took place in Uganda in the past two decades. We then describe how we constructed and validated key variables (refugee presence, development outcomes, control variables) we utilized in the main analysis. These development outcomes include: school access (primary schools and secondary schools), health access index, and road density. Finally, we describe the construction of a supplementary Afrobarometer dataset.

S1.1 Unit of Analysis: Parish-Data Construction

The study's unit of analysis is a parish. Parishes in Uganda are comprised of several nearby villages (median 5 villages per parish with SD=5.5) and they constitute an official administrative unit (local council-2 or LC2, villages are considered the lowest administrative unit, or LC1). In the past two decades, Uganda has experienced substantial proliferation of administrative units (Grossman and Lewis, 2014). According to the National Population and Housing Census Report (2016), the number of parishes increased from 5,238 in 2002 to 7,241 in 2014. As Table S1 makes clear, splits that (mechanically) increase in the number of administrative units took place at all level of local governments.

Table S1: Number of Administrative Units by Census, 1969 2014

	Census Year			
Level of Administrative unit	1969	1991	2002	2014
District	21	38	56	112
County	111	163	163	181
Sub-county	594	884	958	1,382
Parish	3,141	4,636	5,238	7,241

The proliferation of administrative units means that administrative boundaries have changes quite dramatically over the study period. In order to ensure that results across years represent a treatment and not a compositional effect, we had to keep parish boundaries, our unit of analysis, constant across years (2001, 2006, 2011, 2016, 2020). In other words, our first key task was to match and standardize parishes across years and datasets (census data, schools and health facilities data, nightlight data, etc.). We note that this exercise has not been undertaken previously by Ugandan scholars, and as such, we view it as one of the key contributions of our study.

We set our baseline parish boundaries to 2001, based on the mapping exercise of Uganda Bureau of Statistics (UBoS) in preparation for the 2002 census. In other words, 2001 is the benchmark year we selected for all longitudinal empirical analysis for the purpose of boundary consistency. In order to map administrative unit boundaries across years, we used publicly available shapefiles, and more limited crosswalks generated by other scholars. In more details, we considered 2006 parishes to be the same with 2002's and matched directly to 2002's mainly relying on string-based general matching methods (discussed in Section S1.1.1). We generated a 2016-to-2002 crosswalk for mapping 2016 parishes to 2002's (discussed in Section S1.1.2). Another crosswalk which maps 2011 parishes to 2016's (discussed in Section S1.1.3) was also generated in converting 2011 parishes to 2016's first and subsequently to 2002's. With regard to administritive units in 2020, we considered them the same as 2016's.

However, another key challenge stands in the way of making use of these crosswalks. Names and boundaries of different admin levels (district, county, sub-county, parish) are inconsistent across different datasets, even in the same year. For example, some administrative unit names in 2016 are quite different than admin unit names we have in 2016-to-2002 crosswalk where the 2016 admin names come from Uganda 2016 shapefile. Discrepancies are due to either different formatting, minor variations in names used by UBoS or mostly, typos (see Table S2 for examples of frequent inconsistencies). Hence, as a pre-processing step before using crosswalks, we reply on general matching methods again to first match different datasets (i.e. health facility data, school data, etc.) to these crosswalks before they could be used to harmonize the unit of study.

Table S2: Matching problem: examples of inconsistencies

Types	**Examples**
Typo	single character becomes double
	double characters become single
	(ch, c, k), (u, w, y, v), (th, t, s), (r, l)
Minor variation	(west, western), (central, center, centre)
	(town council, T.C., T/C)
	(A parish, A ward, A)
Different formating	- ; _ ; . ; ; /

S1.1.1 General matching methods

String matching we used string matching when identifying non-identical names that describe the same administrative unit across datasets. Instead of using regular expressions, we developed a fuzzy-match algorithm that recognizes matches with one-letter discrepancy for strings less than 6 letters (e.g. *Koboko* VS. *Kobooko*, *Ombachi* VS. *Ombaci*). Strings that have more letters were allowed a discrepancy of 2 letters such as *Bukokho* vs. *Bukhoko*, *Kyegegwa* vs. *Kyegeguua.* We applied fuzzy-match under a fairly strict structured environment, that is, all upper-level administrative names were required to be the same. For example, to increase matching precision, when harmonizing parish level names, we used fuzzy-match to examine parishes under the same district, county, and sub-county.

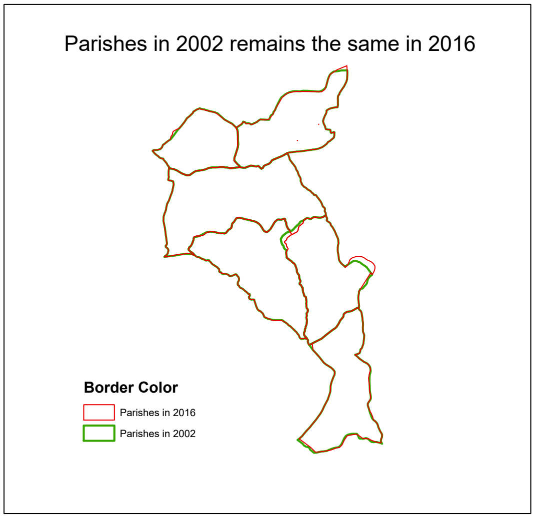
Upper/Lower-level unit tracing is applied when administrative units were aggregated with other units to form a higher-level unit or splinted into different lower-level units. For example, *Kalungu* District in 2011 was a county (also named *Kalungu*) in 2002. Note that villages (LC1s) are also included in this step as they are the lower-level for parishes (LC2s). Applied after string matching, this method first scrutinizes the nearest upper and lower-levels for identical administrative unit names. If failed, all lower level units are compared. If over 50% of the lower units match, the two localities are considered the same no matter how different their names are. For example, Parish A in district D, county C, and sub-county S in 2006 would be matched to parish B in 2002 if all 3 villages of parish A in 2006 appear as villages in parish B in 2002. Note that to apply this rule, parish B needs to also be in district D, county C, and sub-county S.

Keyword matching is mainly needed in matching village names in lower-level unit tracing process and in matching health facility datasets (see Section S1.5.1). We used this method for cases that composed of redundant descriptive or administrative strings like *parish, ward, division, town, primary schools, health centers, facilities* (e.g.*kikyusa (subcounty name) holy cross* vs. *holy cross* vs. *holy cross health center*). These expressions most likely appear in the format of a variety of abbreviations. Fed with a list of possible formats, our keyword matching algorithm was programmed to ignore these descriptive expressions and to only compare the major part (i.e., main administrative names or Health facility names). This method was also used under the restriction that upper-level administrative units must match, again in order to preserve accuracy.
Distance-based matching is only used for matching individual schools (primary and secondary) within each parish, the unit of analysis of our study. This method only applies at the last step, when string-matching, keyword-matching, and manual-matching fail. We construct a distance matrix of all schools in each parish, based on which we pair those that are within 5 km to each other the same facility. This matching strategy hinges on two observable facts: first, it's extremely unlikely that primary or secondary schools would be built within such close distance; second, it's quite common in Uganda that schools changes their name for various reasons such as change of donors. However, this does not apply to health facilities. Health centers at different levels might be fairly close to one another, and they are mostly commonly named after the location whose name does not change often.

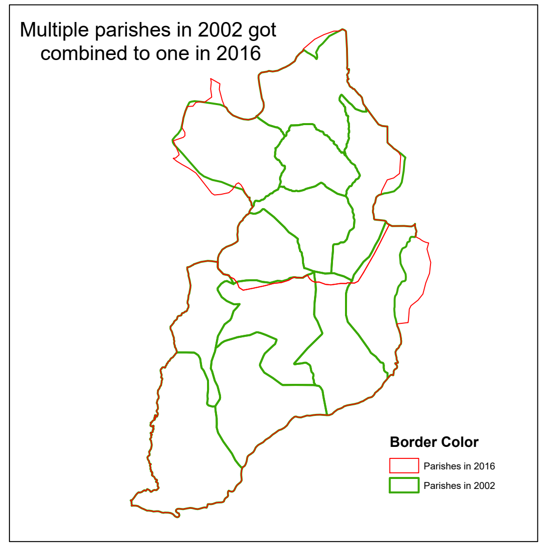
S1.1.2 2016-to-2002 crosswalk

String matching had limited usage when matching 2016 parishes back to 2002. This was the case because parishes in 2016 had substantially redrawn boundaries compared to 2002 boundaries, even for parishes with identical names. Comparing Uganda parish level shapefiles in 2002 and 2016, we found that only 559 parishes in 2002 kept the same boundaries in 2016 (see Figure S1a). By contrast, 1,867 parishes in 2002 were splintered into 2,759 parishes in 2016 (one to multiple, see Figure S1c), and 756 parishes were combined into 719 parishes in 2016 (multiple to one, see Figure S1b). Moreover, the majority of 2002 parishes (2,194) got redistributed rather randomly into 3,464 2016 parishes (see Figure S1d). Again - this haphazard process made string matching impractical.

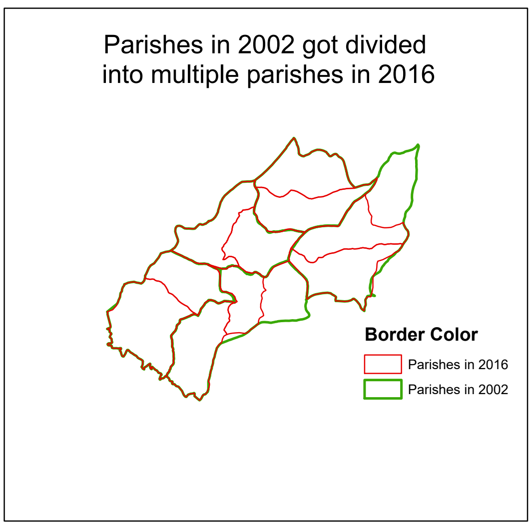
Figure S1: Compare 2002 and 2016 parish shapefiles



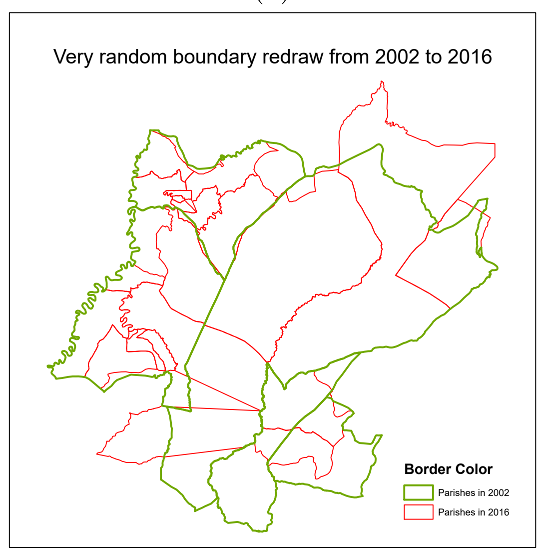
(a)



(b)



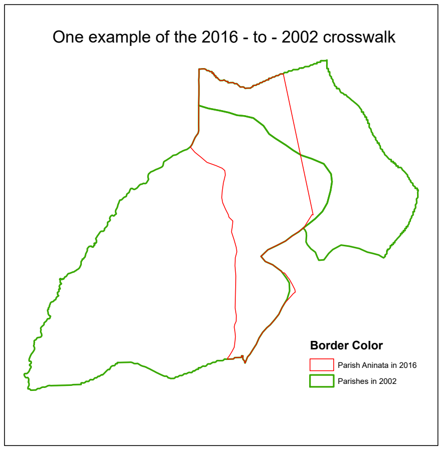
(c)



(d)

Building on shapefiles from 2002 and 2016, we used another approach to map between 2016 and 2002 parishes – an overlapping area method. Specifically, we used the intersection toolkit in ArcGIS and adjusted parameters such that minor misalignment on the boundaries would be disregarded to eliminate potential issues introduced by shapefile digitization errors. Each of the parishes in 2016 was proportionally assigned to 2002 parishes based on the percentage of overlapping areas. Under an additional assumption of evenly distributed population, we were able to allocate census data that are in 2016 parish units to 2002 parishes. Take parish Aninata in 2016 as an example (see Figure S2).

Figure S2: Parish Aninata in 2016



S1.1.3 2011-to-2016 crosswalk

Parish distribution in 2011 is much closer to 2016's than 2002's. Since we already constructed a (relatively) precise 2016-to-2002 crosswalk based on parish overlapping area boundaries as discussed above, generating a 2011-to-2016 crosswalk increased precision in mapping 2011 parishes back to 2002's.

In doing this, we built on an initial mapping generated by Bowles, Larreguy and Woller (2020). These authors mainly used string-matching and lower-level units tracing to identify same polling places across 3 periods of time (before 2013, between the 2013 and 2015 reorganization, 2016) for the purpose of defining the targeting sample of interests that entails fairly specific characteristics in voter registers information. While useful, we identified several problems when implementing this crosswalk for our specific purpose. We corrected these issues with similar matching methods with a fairly stricter and broader searching and matching process. This allowed us to correct and add 492 matches between 2011 and 2016 parishes.

S1.2 Refugee Presence

S1.2.1 Refugee settlements in Uganda

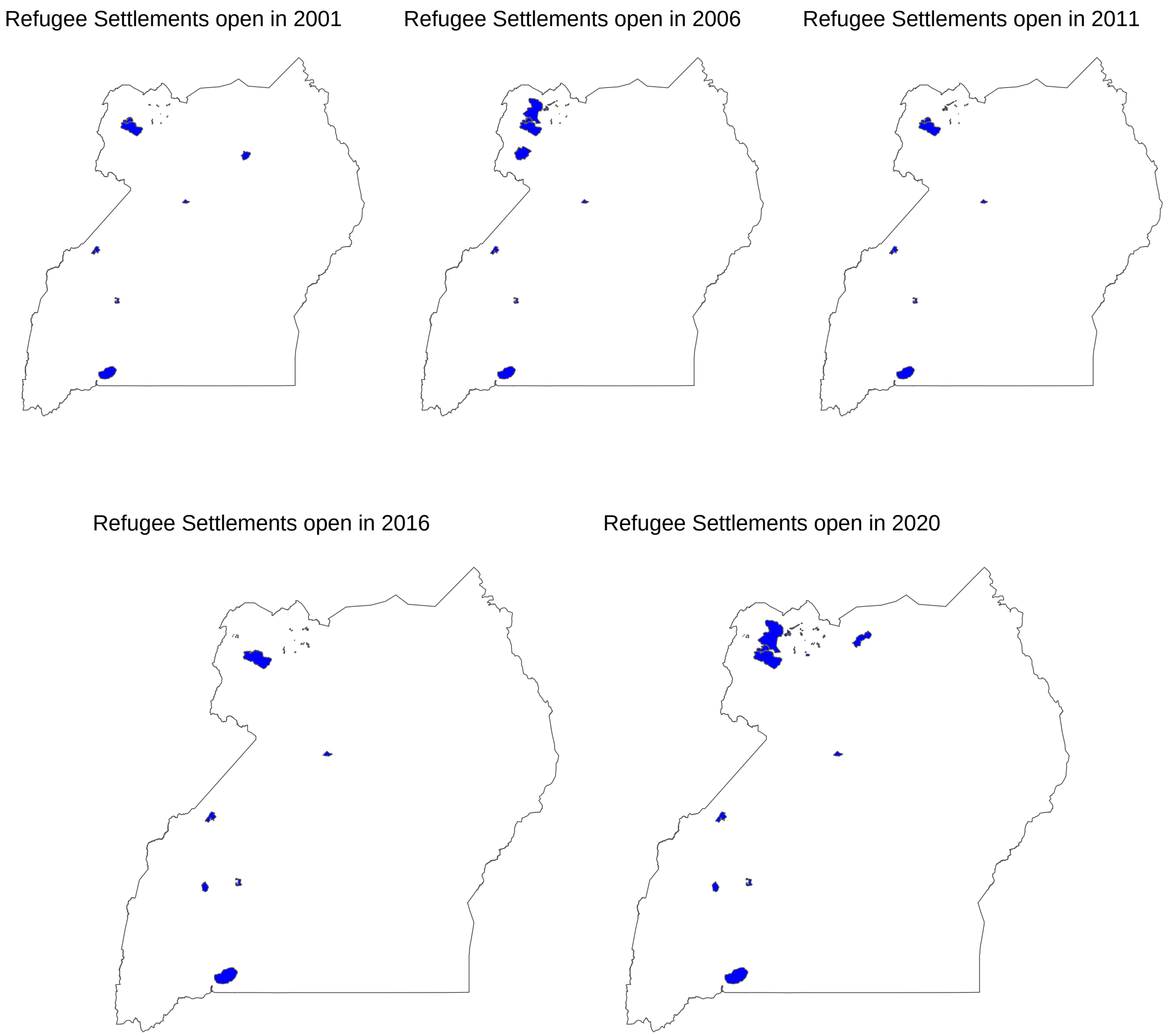


Figure S3: This set of maps shows the refugee settlements in Uganda over the years of our study.

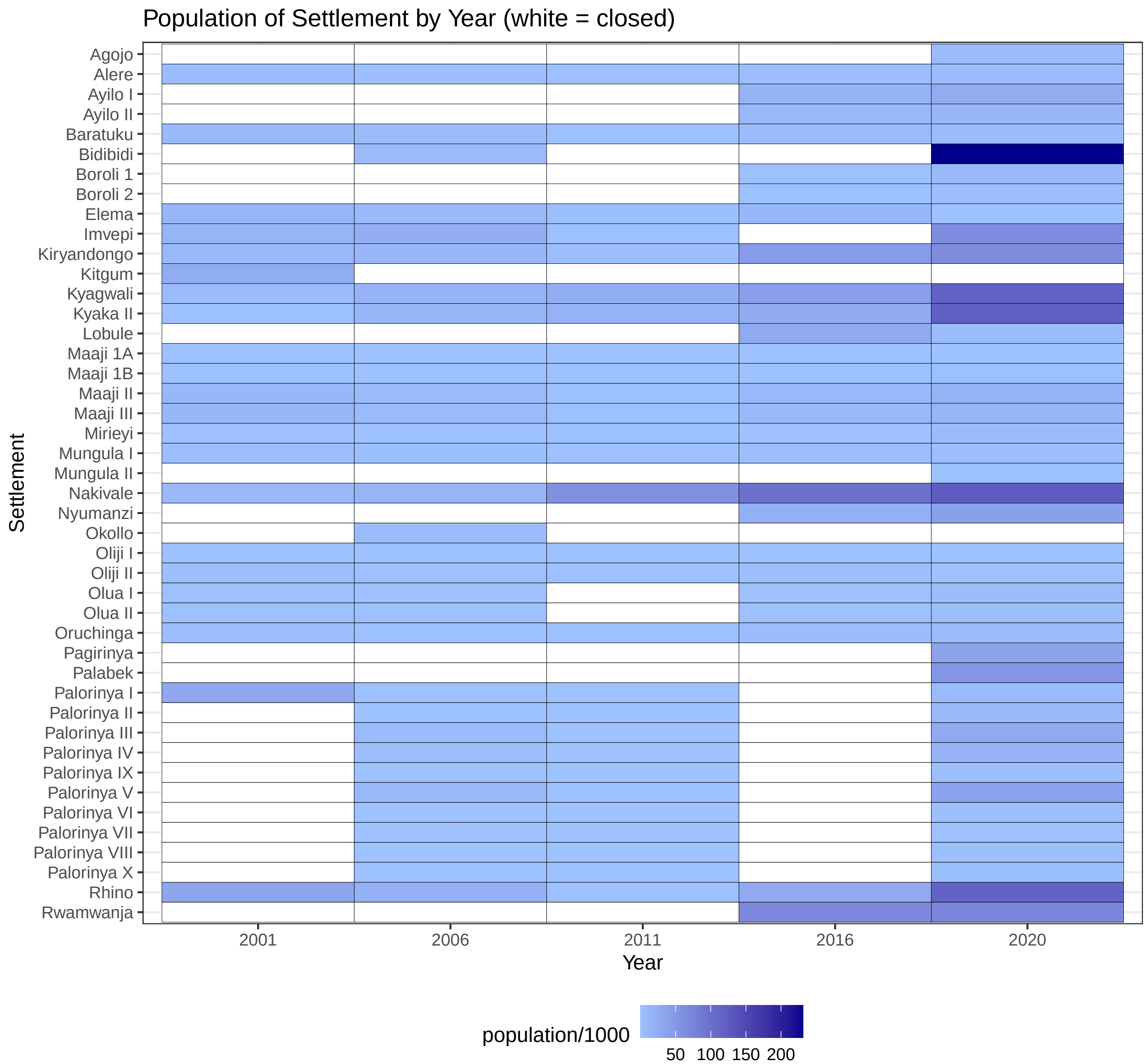


Figure S4: This figure shows the population for each refugee settlement (y-axis) over the years of our study (x-axis).

S1.2.2 Distribution of Presence measures

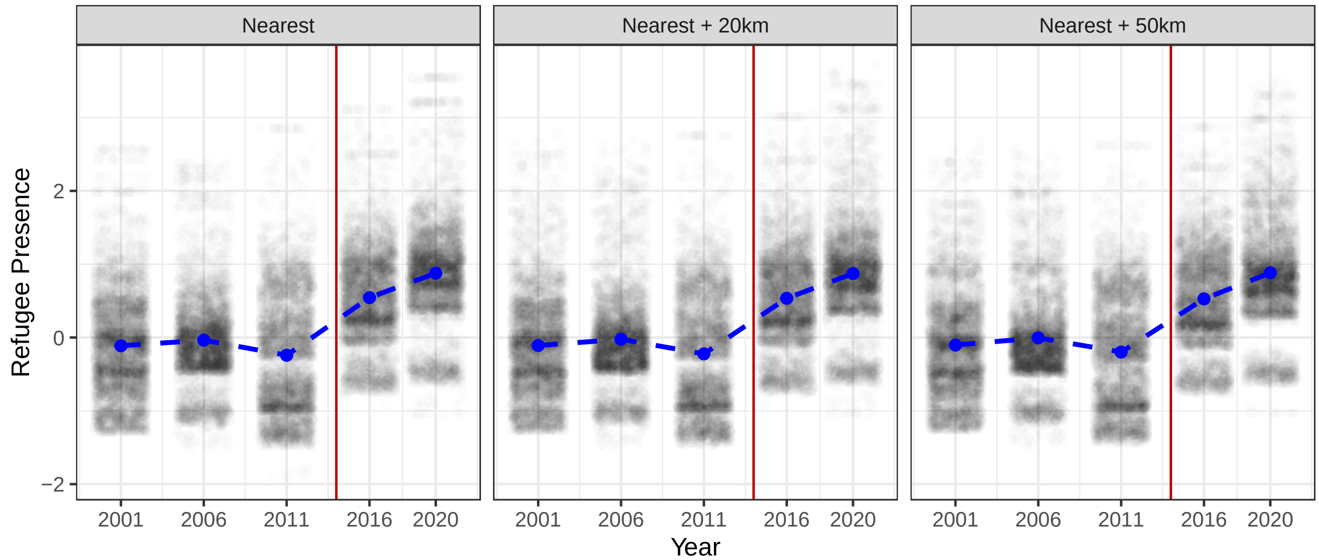


Figure S5: Levels of refugee presence, parishes within 150km. The figure shows that for parishes within 150km of any refugee settlement, our three alternative refugee presence measures all increase after the December 2013 start of the South Sudanese civil war (red line).

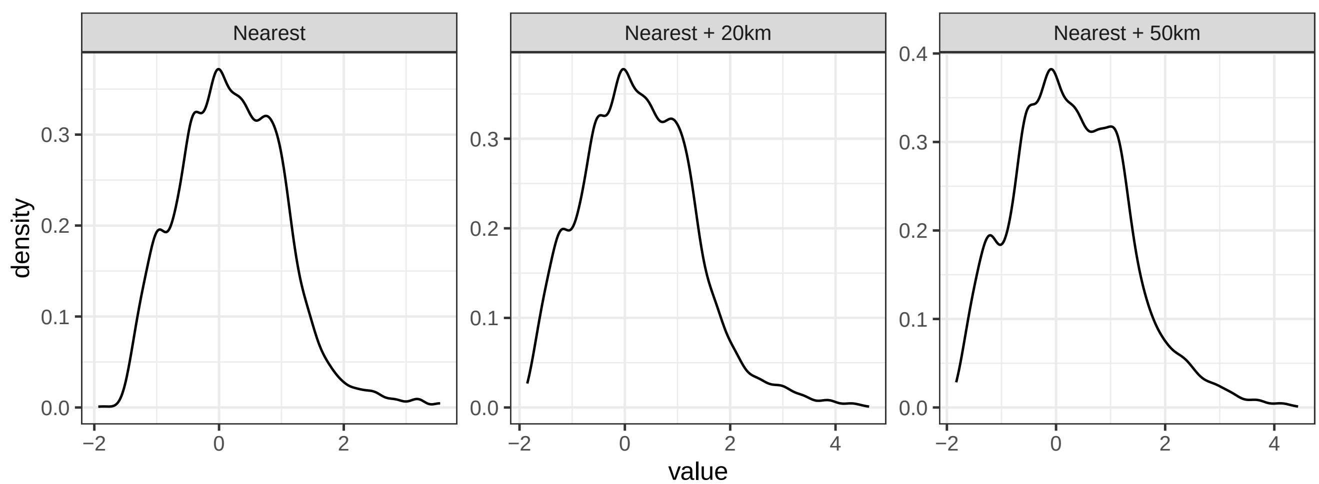


Figure S6: This figure shows the distribution of our refugee presence measures for all parishes within 150km of a refugee settlement.

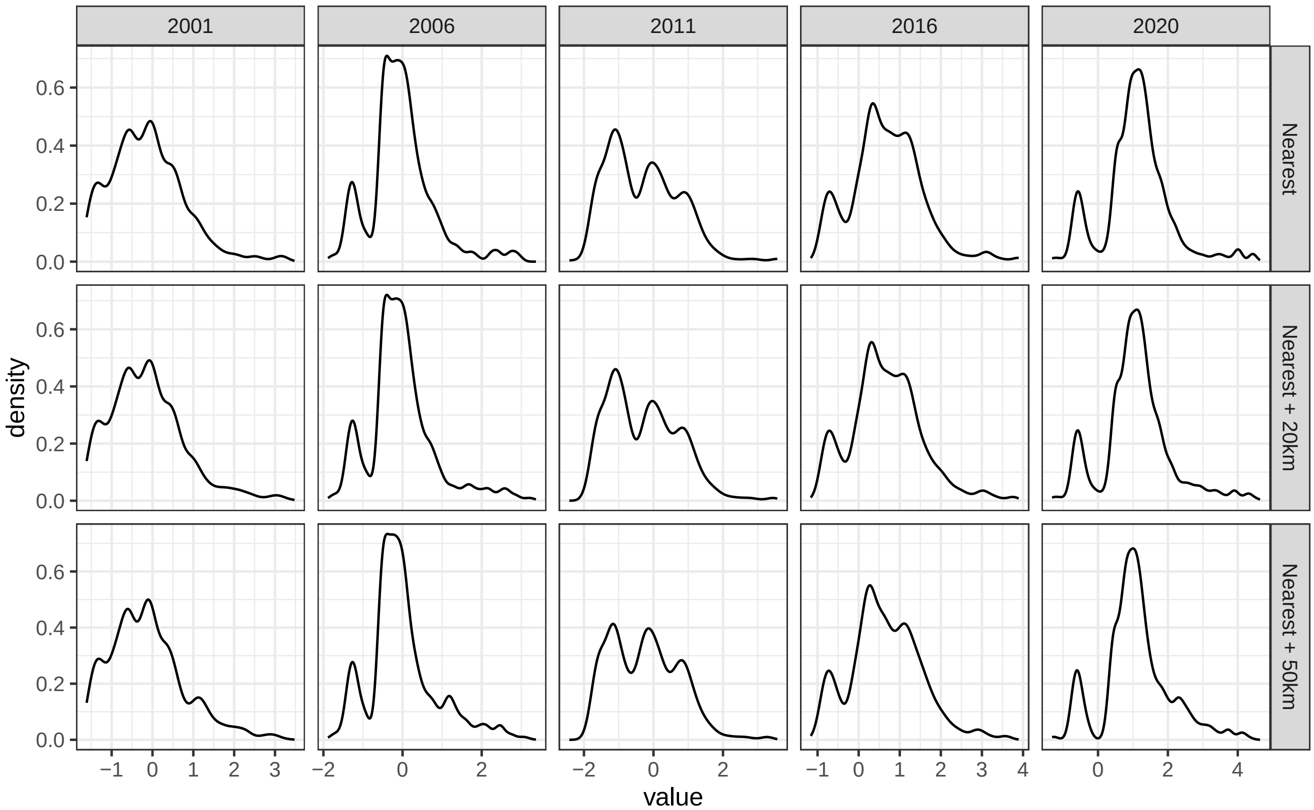


Figure S7: This figure shows the distribution of our refugee presence measures for all parishes within 150km of a refugee settlement, faceted by year.

S1.3 Primary Schools

We generated school access indexes as additional outcome variables. We created a comprehensive primary school list by combining datasets from three sources. The first one is obtained from Uganda Education Management Information Systems (Uganda EMIS) with 19,518 primary schools listed with detailed information including name, ownership, contacts, founding year, and coordinates. This dataset has been verified by crossing-checking with the World Bank's primary school collection. The second dataset is directly shared by the Uganda Bureau of Statistics (UBoS). This dataset extends the time span of the EMIS dataset (before 2017) to date (2020). The last one is collected by our field workers in Uganda which covered 5,277 primary schools with the same attributes. We firstly constructed a comprehensive primary school lists by comparing and merging EMIS dataset and the dataset collected by our field workers.

All three datasets have detailed information about school name, ownership (government-funded versus non government-funded), founding year, location names (from district to parish level), as well as coordinates. We firstly assigned 2002 administrative unit names to all schools based on GPS coordinates, after which we consolidated the three datasets by applying the following general matching methods elaborated in S1.1.1: string matching (applied on school names); keyword matching (applied on school names); distance-based matching. Following such process, we found an overlap of 2,572 primary schools between these the EMIS dataset and our self-collected dataset. Therefore, this list of schools founded prior to 2017 consists of 19,518 EMIS records and 2,705 manually collected schools. Then, we used the up-to-date UBoS dataset as a complementary source to add in newly-founded (after 2017) primary schools. Meanwhile, we update the original list of primary schools founded before 2017 with additional primary schools documented by UBoS. In the end, we have in total 14,136 primary schools existing in 2001, 17,555 in 2006, 23,526 in 2011, 33,132 in 2016, and 34,519 in 2020. Distribution of all schools is mapped with refugee camps (with coral-colored boundaries) (see Figure S8).

Building on this geocoded school list, we further constructed a parish-year cross-sectional dataset by first locating each school in 2002 admin units, then dissecting whether a school existed in 2001, 2006, 2011, 2016, and 2020 separately from its founding year information. Note that we categorized primary schools into two categories based on their ownership: government-funded (named "public") and non-government-funded (named "private"). We also recorded the number of schools in each category for each parish-year in our dataset. We defined the primary school access index to be the number of schools in each parish normalized by parish-level school-aged (6 - 13 years old) population (per thousand). Density plots are presented in Figure S9

Figure S8: Distribution of primary schools.

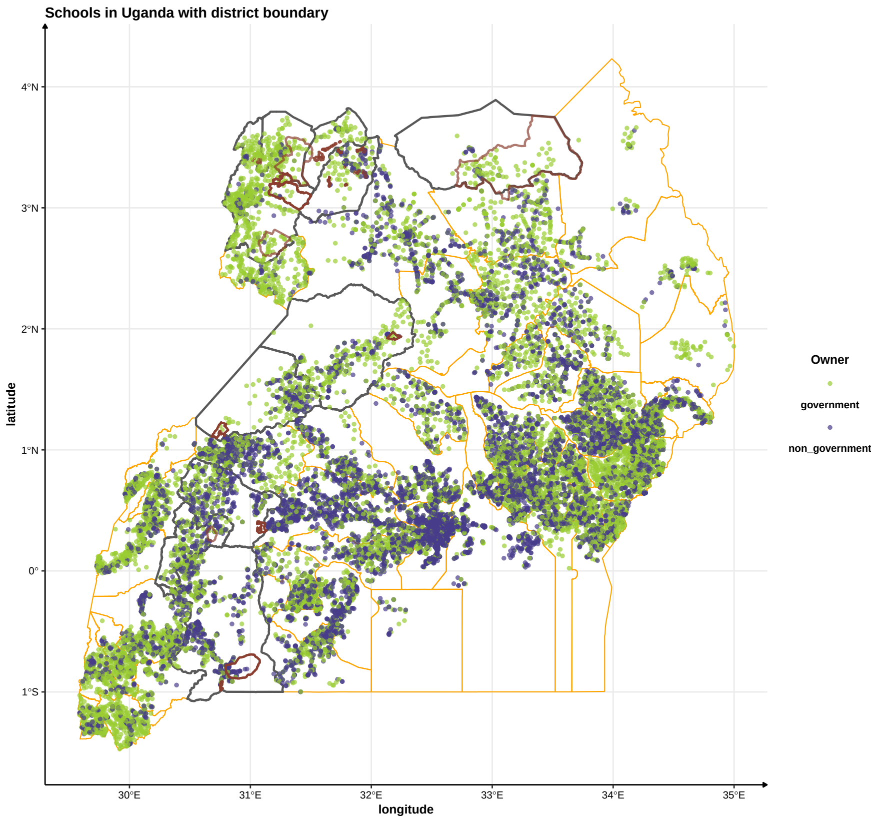
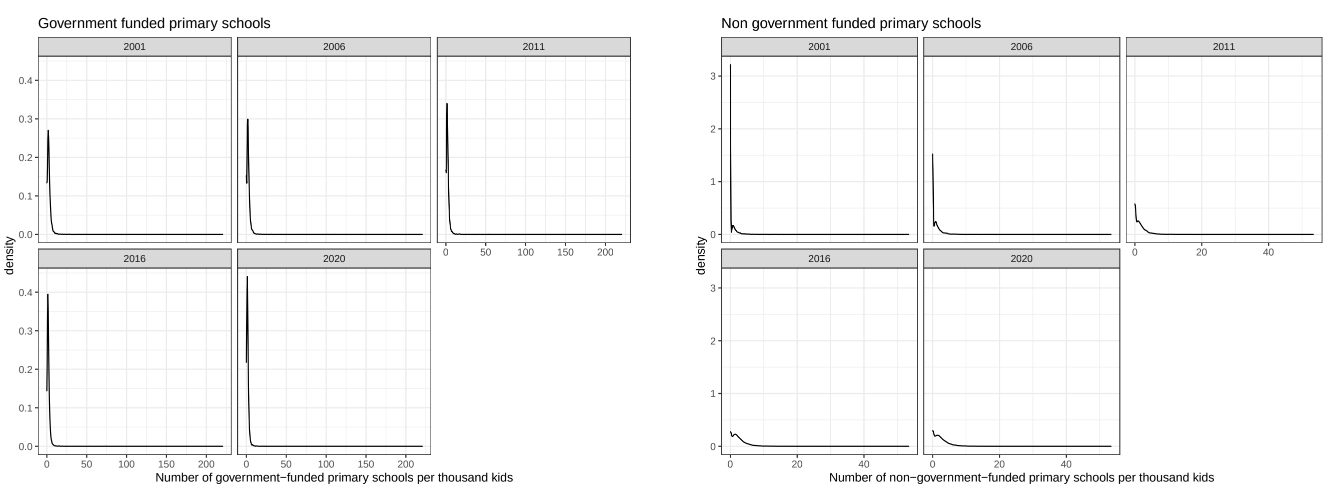


Figure S9: Density plots for primary school access variables

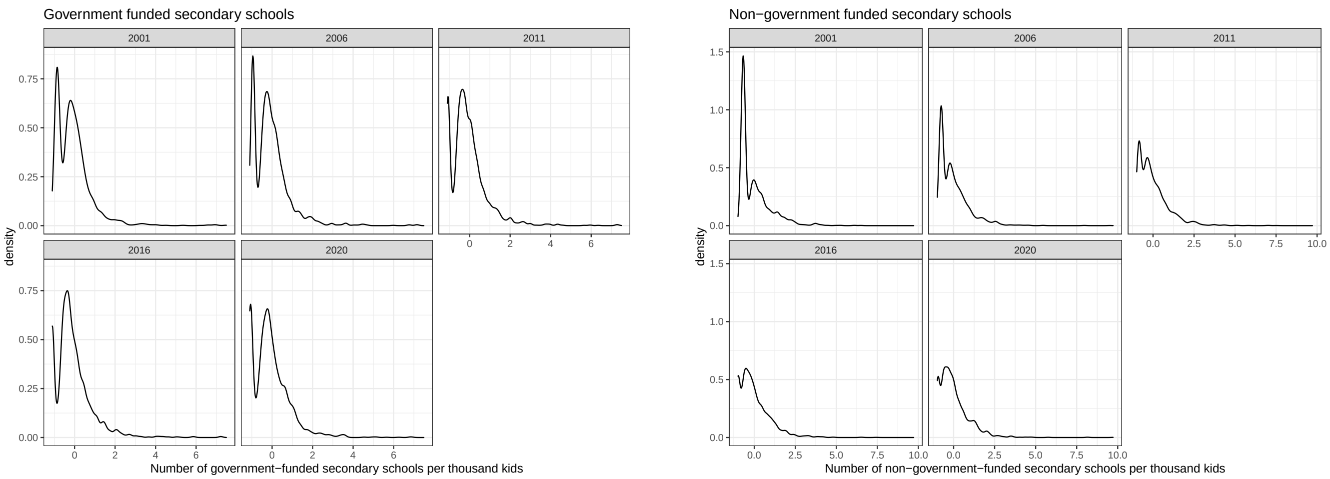


S1.4 Secondary Schools

Secondary schools access is another important outcome variable in our analysis. We exploited a geocoded and verified secondary school data from World Bank, plus a newly-released secondary school list shared by UBoS. Assuming that schools dont close, we constructed a panel dataset with variables indicating if a school existed in 2006, 2011, 2016, or 2020 based on its founding year. Then, one additional index variable is constructed as a proxy of school access following similar process as health access index: the number of schools at the subcounty level normalized by subcounty school-aged (14 - 18 years old) population (per thousand).

Note that this index variable is derived separately for government funded (public), non-government funded (private), and all schools. The school-aged population is calculated from the Worldpop sex-age structured datasets [1]. All variables are standardized to mean 0 and standard deviation 1 at last. Correlation between school access variables and health access index, road density, as well as wealth index range from 0.07 to 0.51.

Figure S10: Density plots for school access variables



(a) (b)

[1] Data can be downloaded at `https://www.worldpop.org/project/categories?id=8`

S1.5 Health Access and Health Utilization (DHS)

Uganda has 5 levels of public health facilities: HC I (clinic), HC II, HC III, HC IV, HC V (Hospital) (from the smallest to largest in scale). Each serves (at least in theory) a different administrative level (recall, Uganda has five levels of administrative units from the village (LC1) to the district (LC5)). Importantly, the different levels of health facility correspond to different levels of capacity, equipment, staff size, and available services. The most basic unit is the health clinic-1 (HC1), which operates at the village level and is usually staffed by community health workers with no formal nursing training; HC2s are deigned to serve several nearby villages; HC3s are designed to serve a subcounty, staffed by nurses, and offer basic laboratory testing services; HC4 are hospitals with limited range of services and finally, HC5s are regional hospitals offering a relatively broad set of services, including surgery capabilities (MOH, 2016).

A valid health access index should be able to capture three dimensions: (a) *distance* to the nearest health facility; (b) *population served* by the facility, and (c) the *range of services* offered by the facility. In general, a parish has better health access when people travel shorter distances to facilities serving fewer people while also offering more services.

S1.5.1 Health facility dataset construction

Data source: A complete national Master Health Facility list (MFL) for Uganda is not publicly available (Mpango and Nabukenya, 2019). We managed to obtain drafts of Uganda's 2020, 2017 and 2012 Master facility list from the Ministry of Health (MoH), and the 2006 health facility list from the Uganda Bureau of Statistics (UBoS).

Data attributes: There are 3,339 health facilities listed in the 2006 dataset, 5,410 in 2012, 6,248 in 2016, and 7,813 in 2020. All datasets contains name, level, owner, authority of the facility and the district, county (except for 2016 dataset), subcounty, and parish name. Around 57% facilities in 2012 dataset, 64% in 2016, and all facilities in the 2020 list additionally have exact coordinates documented.

Across data sources we found both duplicate records and large discrepancies in basic facility information such as facility name, location, type, level, ownership, and other attributes. Among them, facility name and administrative unit names are most consequential, because we rely on them to locate and identify each facility, especially those that do not have available GIS coordinates. The inconsistencies exist not only across years but also within datasets. In addition, for some facilities, recorded coordinates lead to a location faraway from the district, county, subcounty, and parish that were separately documented for this facility. These challenges were addressed in the following merging and cross-checking process.

Merge across years: We merged the four datasets and constructed a consolidated comprehensive facility list with the name, level, owner, authority of each facility, whether it exist in a certain year (2006, 2012, 2016, or 2020), administrative unit names and coordinates for the purpose of cross-checking and cross-referencing. The facility list of 2012 was used as a base data with 2006, 2016, and 2020 facilities matched to it separately. Unlike merging election data, we did not use crosswalks here because for half of 2012 data, 2/3 of 2016 data, and all 2020 data, we can directly use coordinates to locate facilities in 2002 parishes using GIS. After assigning 2002 administrative unit names to those that have coordinates, we cleaned and matched 2006 administrative names and the rest one third of 2016 facilities without coordinates level by level, as well as facilities names with 2012 dataset using general matching methods and 2016-to-2002 crosswalk. Although the crosswalk works the best in calculating percentages under the assumption of even distribution, we used it to evenly distribute 2016 facilities to each of the matched 2002 parishes, considering that it is plausible to think that residents have an equal ability to "access" that facility. 3,132 facilities in 2006 dataset were also found existing in 2012 data, 3,799 exists in both 2016 and 2012, and only 4,220 out of 7,813 facilities in the 2020 master lists were built after 2016.

Cross-check and supplement geocoding: Using 2012 as the base year, we combined the above two matched and merged datasets (2006-merged-to-2012 dataset and 2016-merged-to-2012 dataset) together with the 2020 dataset, with in total around 9,500 facilities listed. The majority now has coordinates that could be used to assign a 2002 parish. For the rest that only have administrative unit names available, we used reverse-geocoding with Google Maps API to search for coordinates. By feeding in health facility names including "Health center, Uganda" string, Google Maps were able to return coordinates for a small amount of facilities. Results were double-checked by comparing provided admin unit names with parishes the coordinates lead to. Lastly, we manually searched for about 200 HC III / IV / V's for which we did not have exact coordinates. We were able to successfully identify 146 additional health facilities which we verified by cross-checking with provided district / subcounty / parish names.

In the end, we successfully identified the location of around 7,100 facilities using coordinates, and located 2,400 additional facilities in terms of their 2002 parish using administrative names. Those facilities were assigned parish centroid coordinates.

S1.5.2 Health Access index construction

Based on the comprehensive geocoded health facility list containing 2006, 2011, 2016, and 2020 information, we constructed several parish-year variables for the 5 different levels of HC. We constructed three variables for HC I: the number of HC I; an indicator of whether there is at least one HC I (YES = 1, NO = 0); the number of HC I normalized by population at subcounty level. For HC II, we also normalized the number of HC II's in each subcounty by the population as an estimation of access. For HC III / IV / V where we have the most accurate coordinates, we constructed a measure of the shortest **Distance** from each parish's centroid to a facility and a measure of **Crowdedness** (the population served by the closet health center). Crowdedness is defined as the sum of population of all parishes that are closest to a given facility.

S1.5.3 Health utilization

We used the Uganda Demographic and Health Survey (DHS) to measure health services utilization, which is available for the three years of health facility data, 2006, 2011, and 2016. We mainly used DHS's Kids Recode (KR) data and Geographic Data (GE) which is used to geo-locate each respondent. The unit of analysis of KR is each child under 5 years old born to an interviewed reproductive aged women (15 - 49). The KR dataset contains both the information related to the child such as immunization, nutrition, health condition, and information for the mother such as pre/post-natal care for each pregnancy. **Further, we also incorporated household possession of malaria-prevention nets from the Household Records (HR)** as additional variables appended to each kid. In total, we have 7,634 observations included in DHS 2006, 7,796 in 2011, and 15,277 in 2016. **We examined all possible statistics provided in DHS Guide to Statistics, selected questions that indicate people's utilization of health services and**

standardized them across 2006, 2011, 2016 DHS data because questions and variables have been modified across years. After testing correlation and cronbach's alpha, we categorized utilization variables into women's utilization and child's utilization (Magadi, Agwanda and Obare, 2007; Dimbuene et al., 2018; Gyimah, Takyi and Addai, 2006; Yaya et al., 2020; Adhikari, 2016):

- **women's utilization: including maternal health, possesion of mosquito nets and Insecticide-Treated Net (ITN);**
- **child's utilization: including child's vaccination, deworming, and iron supplements provision.**

For each kid in KR data, we calculated the ultimate utilization index as the simple average across all variables that include both its own utilization variables and its mother's maternal variables. Details about recoding each variables are explained in the following.

For women's utilization, we recoded and generated the following variables: (1) **Tetanus injection before pregnancy:** takes the value "1" if the woman has received a tetanus injection before pregnancy; (2) **Tetanus injection during pregnancy:** takes the value "1" if the woman has received a tetanus injection during pregnancy; (3) **Adequate ANC visit:** takes the value of 1 if at least four antenatal (ANC) visits were made; (4) **ANC provider:** and (5) **Delivery assistance:** we consider the following categories "professional personnel" and assign 1 to this dummy variable: doctor, nurse / midwife, medical assistant/clinical officer, nursing aide / assistant; (6) **Places of ANC:** (7) **Places of Delivery:** Places of ANC and Places of Delivery variable are coded 1 if the ANC or delivery happened at government / private hospital, health center, and clinics, or other public / private sectors compared to residential houses; (8) **Household possession of mosquito net:** takes the value of 1 if the household have at least one mosquito net; (9) **Household possession of Insecticide-Treated Net:** takes the value of 1 if the household have at least one insecticide-treated net.

For child's utilization, our primary indicator is whether a child was vaccination, received deworming medicine, and iron supplements. With regard to vaccination, we consider "complete_vaccination" for children receiving 1 BCG, 3 DTP, 3 POLIO, 1 MCV before the age of 3. Following this, we focused on the following 4 individual indicators. (1) **BCG immunization:** takes a value of 1 if the child receives 1 dose of BCG; (2) **Complete DPT immunization:** takes a value of 1 if the child receives 3 doses of DPT; (3) **Complete POLIO immunization:** takes a value of 1 if the child receives 3 doseS of POLIO; (4) **Complete MCV immunization:** takes a value of 1 if the child receives 1 dose of MCV. (5) **Dewoming:** takes 1 if the child received deworming medication; (6) **Iron supplements:** takes 1 if the child recieved iron supplements.

S1.6 Road Density

We calculated the Speed-Weighted-Road-Density in capturing and describing road quality for each parish. It is defined as the ratio of total length of roads (km) to the area (km^2), weighted by the speed limit of each type (see Equation S1).

$$Weighted\_Road\_Density_k = \frac{\sum_{i=1}^{i=6} Length_i * Velocity_i}{\sum_{i=i}^{i=6} Velocity_i} \div Area_k \tag{S1}$$

with k representing each parish, i is the road class from Highway to Trail / Track.

To calculate road density for each parish, We first extracted the length of roads in each road class that falls within each parish using geographic information systems (GIS), categorized by road classes (See Figure S11). We mainly used three shapefiles of Uganda road network. First one is the most recently updated (in 2020) road network shapefile extracted from OpenStreetMap by the Humanitarian OpenStreetMap Team. [2]. The second road map is constructed by World Food Programme using OpenStreetMap and last updated in 2017, post-treatment period in our study.[3] The last one is from Global Roads Open Access Data Set gathered by NASA Socioeconomic Data and Applications Center (SEDAC), covering road information up to 2010 (Center For International Earth Science Information Network-CIESIN-Columbia University and Information Technology Outreach Services-ITOS-University Of Georgia, 2013). [4]

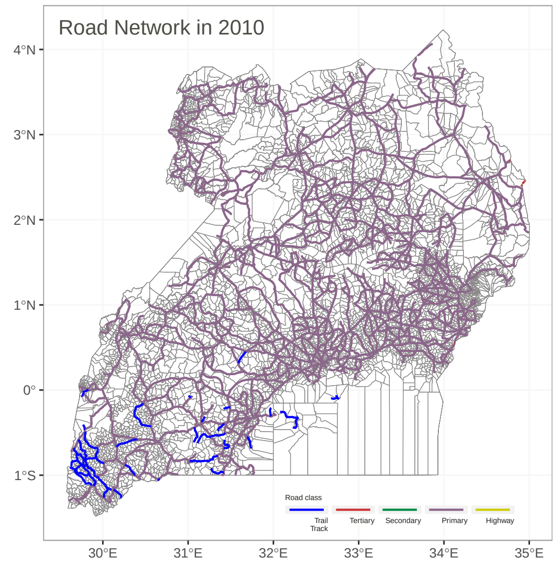
---

[2]Shapefile can be downloaded at the Humanitarian Data Exchange (HDX) website `https://data.humdata.org/dataset/hotosm_uga_roads`

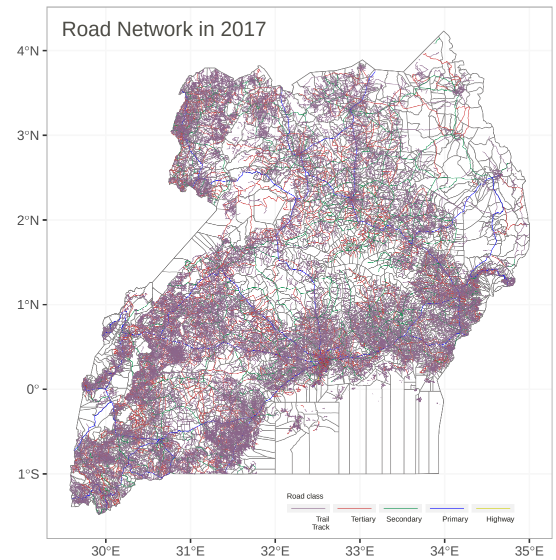
[3]Shapefile can be downloaded at WFP website `https://geonode.wfp.org/layers/geonode:uga_trs_roads_osm`

[4]Shapefile can be downloaded at NASA SEDAC website `https://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1`

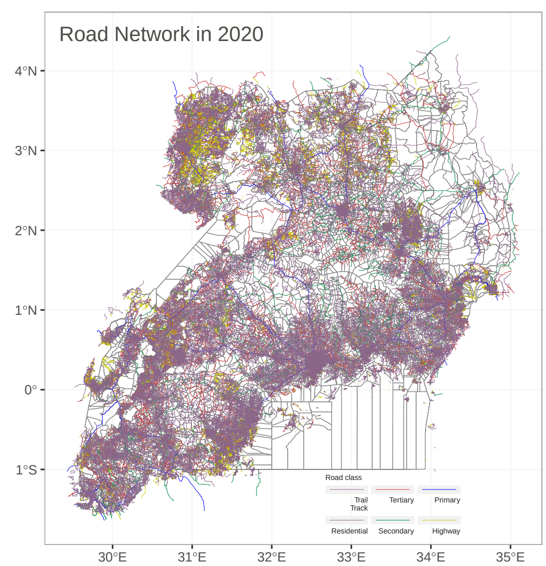
Figure S11: Road Network in Uganda



(a) Road network in 2010



(b) Road network in 2017



(c) Road network in 2020

There are mainly 6 types of road class in Uganda, from the best to the worst: Highway; Primary; Secondary; Tertiary; Residential or Local; Trail or Track. Very few pathways, private roads, and unspecified roads that are in the shapefiles are excluded in our calculation. The speed limit for each road class were generated by Carl Müller-Crepon (2021) from Michelin website `www.viamichelin.com`. We recoded the speed of highways to that of the second category to maintain the rank order.

Further, we tested the Spearman's correlation between road density and our other measures: wealth index, nighttime light density, health access index.

Table S3: Correlation

	Road Density		
	2010	2017	2020
Secondary School	0.191***	0.349***	0.256***
Primary School	0.034***	0.182***	0.110***
Health Index	0.177***	0.486***	0.524***

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

S1.7 Afrobarometer Survey Data

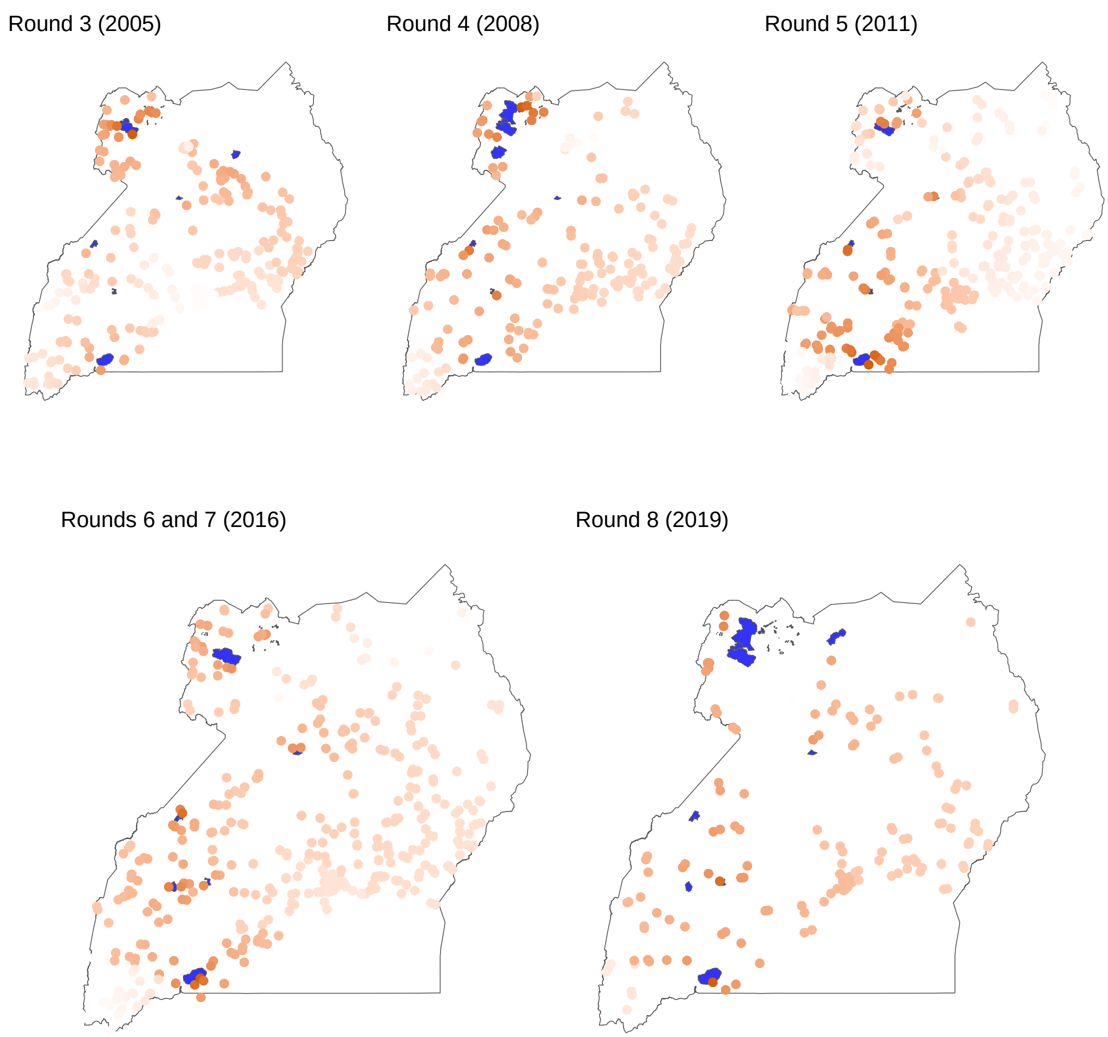


Figure S12: This set of maps shows across our study years, the locations of refugee settlements (blue) and Afrobarometer respondents (orange) shaded by their refugee presence levels to the settlements.

We further explored nationally representative Afrobarometer (AB) survey data to analyze support for the president, the NRM party, the government; attitudes towards migrants and migration policy; and perceptions of insecurity. We utilized geocoded AB data from Round 3 (2005, N = 2391), Round 4 (2008, N = 2421), Round 5 (2011, N = 2380), Round 6 and 7 (2015 and 2016, N = 3456), and Round 8 (2019, N = 1199). Figure S12 shows the geographic distribution of the respondents (orange) by round, shaded by their refugee presence levels to the refugee settlements (blue).

Questions and response coding vary across years. We selected pertinent questions (as displayed in Table S4) and standardized the responses across different rounds. For migrants as neighbors, migrants can move freely, born to a non-Ugandan, and able to naturalize, we rescaled the variables so that higher values are more pro-migrant and supportive of open citizenship. For felt unsafe in community, feared crime, higher values means more fear.

Table S4: Afrobarometer variables

Categories	**Variables**
Administrative	round, year, respondent number, region, district
Demographic	urban or rural, ethnicity, religion, gender, race, age, occupation, education, language
Wealth	living condition, living condition compared to others, radio, television, motorcycle/vehicle, computer, mobile phone, enough food/water/cooking fuel/cash income/school expenses, have internet access, have phone services, roof material, shelter type
Govt Performance	improve economy/poor people's condition/health services/education services/ water condition/food supply/road condition/electricity supply, create job opportunities, stabilize price, protect forest and river, empower women, decrease inequality/crime/corruption/hiv-aids/conflicts
Migration Attitudes	migrants as neighbors, migrants can move freely, born to a non-Ugandan, able to naturalize
Insecurity	felt unsafe in community, fear crime

All outcome variables are standardized with mean 0 and standard deviation being 1. We also generated a wealth index and a government performance index using categorical principle component analysis. Variables that were included in indexes are slightly different for the four rounds because some variables are not available in certain rounds.

S1.8 ACLED Violence data

We collected violent events data in one year prior to our study years (2001, 2006, 2011, 2016, 2020) from the Armed Conflict Location & Event Data Project (ACLED). These events were categorized in major groups including riots, battles, protests, and violence and further in more detailed sub-event groups. We assigned 2002 parish ID to each violent event using GIS and summarized the number of events for each main event group. Further, we generated the dichotomous outcome variable *any_violent_event* which takes 1 if any of the following 5 main events happened in that parish-year:violence against civilians, riot, attack, mob violence, violent event.

S2 Robustness Checks

This section shows the regression tables for the main results with alternative specifications.

S2.1 Alternative Refugee Presence Measures

In this section, we use Nearest and Nearest + 50km refugee presence measures.

	Public Pri	Public Sec	Roads	PG Index	HC 2	HC 3	HC 4	HC 5	Health Util
Baseline Presence	-0.046	-0.017^{***}	0.033^{*}	-0.029^{***}	0.029	-0.177^{***}	-0.279^{***}	0.085^{***}	-0.125^{***}
	(0.053)	(0.006)	(0.020)	(0.005)	(0.035)	(0.035)	(0.033)	(0.024)	(0.015)
Presence x 2006	0.118^{**}	0.015^{***}		0.025^{***}	-0.207^{***}	0.116^{***}	0.264^{***}	-0.141^{***}	
	(0.046)	(0.005)		(0.005)	(0.032)	(0.035)	(0.031)	(0.021)	
Presence x 2011	0.108^{**}	0.014^{**}		0.036^{***}	-0.049^{*}	0.135^{***}	0.279^{***}	-0.105^{***}	0.099^{***}
	(0.048)	(0.006)		(0.005)	(0.028)	(0.030)	(0.030)	(0.021)	(0.020)
Presence x 2016	0.159^{***}	0.018^{**}	0.058^{***}	0.033^{***}	-0.150^{***}	0.186^{***}	0.256^{***}	-0.049^{**}	0.050^{***}
	(0.057)	(0.007)	(0.014)	(0.005)	(0.029)	(0.030)	(0.031)	(0.023)	(0.018)
Presence x 2020	0.275^{***}	0.034^{***}	0.096^{***}	0.020^{***}					
	(0.077)	(0.009)	(0.015)	(0.005)					
Diff 2016-2011	0.051	0.004		-0.003	-0.102	0.051	-0.024	0.056	-0.049
SE Diff 2016-2011	0.018	0.002		0.001	0.009	0.011	0.008	0.010	0.016
Diff 2020-2011	0.167	0.020		-0.016					
SE Diff 2020-2011	0.046	0.004		0.003					
Diff 2020-2016	0.116	0.016	0.038	-0.014	-0.014	-0.014	-0.014	-0.014	
SE Diff 2020-2016	0.036	0.003	0.012	0.003	0.003	0.003	0.003	0.003	
Presence Measure	Nearest	Nearest	Nearest	Nearest	Nearest	Nearest	Nearest	Nearest	Nearest
Sample Distance (km)	150	150	150	150	150	150	150	150	150
Controls x Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	18185	18185	10911	18185	14548	14548	14548	14548	17982
R^2 (full model)	0.871	0.845	0.727	0.865	0.545	0.662	0.760	0.861	0.190
R^2 (proj model)	0.038	0.055	0.104	0.160	0.108	0.044	0.112	0.053	0.056
Num. groups: parish_id	3637	3637	3637	3637	3637	3637	3637	3637	
Num. groups: region_year	20	20	12	20	16	16	16	16	
Num. groups: region									4
Num. groups: year									3

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table S5: Regression table for Public Goods outcomes: For parishes within 150km, OLS models interacting Nearest refugee presence (time-varying) with year, interacting controls with year, and including parish FE and region-year FE, with standard errors clustered at the parish level.

	Public Pri	Public Sec	Roads	PG Index	HC 2	HC 3	HC 4	HC 5	Health Util
Baseline Presence	-0.072	-0.014^{*}	0.052^{**}	-0.028^{***}	0.031	-0.179^{***}	-0.304^{***}	0.094^{***}	-0.118^{***}
	(0.068)	(0.007)	(0.024)	(0.006)	(0.038)	(0.038)	(0.036)	(0.027)	(0.014)
Presence x 2006	0.180^{***}	0.013^{***}		0.030^{***}	-0.203^{***}	0.122^{***}	0.302^{***}	-0.138^{***}	
	(0.059)	(0.005)		(0.005)	(0.034)	(0.035)	(0.033)	(0.023)	
Presence x 2011	0.187^{***}	0.013^{*}		0.038^{***}	-0.067^{**}	0.142^{***}	0.304^{***}	-0.107^{***}	0.104^{***}
	(0.061)	(0.007)		(0.005)	(0.029)	(0.032)	(0.032)	(0.022)	(0.020)
Presence x 2016	0.259^{***}	0.018^{**}	0.062^{***}	0.036^{***}	-0.157^{***}	0.188^{***}	0.274^{***}	-0.065^{***}	0.043^{**}
	(0.073)	(0.008)	(0.015)	(0.005)	(0.030)	(0.031)	(0.033)	(0.024)	(0.017)
Presence x 2020	0.388^{***}	0.037^{***}	0.080^{***}	0.020^{***}					
	(0.090)	(0.009)	(0.016)	(0.005)					
Diff 2016-2011	0.071	0.006		-0.002	-0.091	0.046	-0.030	0.042	-0.061
SE Diff 2016-2011	0.020	0.003		0.001	0.009	0.011	0.008	0.010	0.017
Diff 2020-2011	0.200	0.025		-0.018					
SE Diff 2020-2011	0.048	0.005		0.003					
Diff 2020-2016	0.129	0.019	0.018	-0.016	-0.016	-0.016	-0.016	-0.016	
SE Diff 2020-2016	0.035	0.003	0.013	0.003	0.003	0.003	0.003	0.003	
Presence Measure	Nearest+50	Nearest+50	Nearest+50	Nearest+50	Nearest+50	Nearest+50	Nearest+50	Nearest+50	Nearest+50
Sample Distance (km)	150	150	150	150	150	150	150	150	150
Controls x Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	18185	18185	10911	18185	14548	14548	14548	14548	17982
R^2 (full model)	0.871	0.845	0.726	0.865	0.545	0.662	0.761	0.861	0.190
R^2 (proj model)	0.040	0.056	0.102	0.161	0.107	0.043	0.115	0.052	0.056
Num. groups: parish_id	3637	3637	3637	3637	3637	3637	3637	3637	
Num. groups: region_year	20	20	12	20	16	16	16	16	
Num. groups: region									4
Num. groups: year									3

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table S6: Regression table for Public Goods outcomes: For parishes within 150km, OLS models interacting Nearest + 50km refugee presence (time-varying) with year, interacting controls with year, and including parish FE and region-year FE, with standard errors clustered at the parish level.

	Neighbors	Move Freely	Felt Unsafe	Feared Crime	Born Non-Ugandan	Naturalize
Baseline Presence	0.077^{***}	-0.014	0.055^{**}	-0.010	-0.106^{***}	0.091^{***}
	(0.026)	(0.033)	(0.024)	(0.032)	(0.032)	(0.032)
Presence x 2006				-0.061		
				(0.048)		
Presence x 2011				0.064		
				(0.044)		
Presence x 2016	0.006		-0.045	0.208^{***}		
	(0.052)		(0.048)	(0.052)		
Presence x 2020	-0.014	0.037	-0.118^{**}	-0.121^{**}		
	(0.053)	(0.055)	(0.049)	(0.053)		
Diff 2016-2011				0.144		
SE Diff 2016-2011				0.051		
Diff 2020-2011				-0.185		
SE Diff 2020-2011				0.052		
Diff 2020-2016	-0.020		-0.073	-0.329		
SE Diff 2020-2016	0.062		0.057	0.060		
Presence Measure	Nearest	Nearest	Nearest	Nearest	Nearest	Nearest
Sample Distance (km)	150	150	150	150	150	150
Controls x Year	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	2830	1994	2824	5906	1409	1389
R^2 (full model)	0.061	0.025	0.020	0.058	0.029	0.080
R^2 (proj model)	0.028	0.014	0.018	0.052	0.022	0.031
Num. groups: region	4	4	4	4	4	4
Num. groups: year	3	2	3	5		

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table S7: Regression table for Afrobarometer outcomes: For respondents within 150km, OLS models interacting Nearest refugee presence (time-varying) with year, interacting controls with year, and including region and year FE.

	Neighbors	Move Freely	Felt Unsafe	Feared Crime	Born Non-Ugandan	Naturalize
Baseline Presence	0.088^{***}	-0.035	0.045^{*}	-0.007	-0.089^{***}	0.124^{***}
	(0.027)	(0.033)	(0.025)	(0.030)	(0.033)	(0.033)
Presence x 2006				-0.083^{*}		
				(0.042)		
Presence x 2011				0.076^{*}		
				(0.043)		
Presence x 2016	0.009		-0.103^{**}	0.163^{***}		
	(0.052)		(0.048)	(0.050)		
Presence x 2020	-0.022	0.058	-0.101^{**}	-0.129^{**}		
	(0.052)	(0.055)	(0.048)	(0.051)		
Diff 2016-2011				0.087		
SE Diff 2016-2011				0.052		
Diff 2020-2011				-0.205		
SE Diff 2020-2011				0.052		
Diff 2020-2016	-0.031		0.002	-0.293		
SE Diff 2020-2016	0.062		0.057	0.059		
Presence Measure	Nearest+50	Nearest+50	Nearest+50	Nearest+50	Nearest+50	Nearest+50
Sample Distance (km)	150	150	150	150	150	150
Controls x Year	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	2830	1994	2824	5906	1409	1389
R^2 (full model)	0.062	0.025	0.020	0.058	0.026	0.084
R^2 (proj model)	0.029	0.015	0.018	0.052	0.019	0.035
Num. groups: region	4	4	4	4	4	4
Num. groups: year	3	2	3	5		

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table S8: Regression table for Afrobarometer outcomes: For respondents within 150km, OLS models interacting Nearest + 50km refugee presence (time-varying) with year, interacting controls with year, and including region and year FE.

S2.2 Alternative Sample Radii

For the alternative sample radii, we show results using parishes within 100km from a refugee settlement at baseline, 200km, and all parishes.

	Public Pri	Public Sec	Roads	PG Index	HC 2	HC 3	HC 4	HC 5	Health Util
Baseline Presence	-0.021	-0.017^{**}	0.046^{**}	-0.010	-0.020	-0.127^{***}	-0.108^{***}	0.117^{***}	-0.100^{***}
	(0.057)	(0.008)	(0.023)	(0.006)	(0.039)	(0.046)	(0.037)	(0.030)	(0.017)
Presence x 2006	0.106^{***}	0.013^{**}		0.005	-0.155^{***}	0.075^{*}	0.097^{***}	-0.158^{***}	
	(0.037)	(0.005)		(0.006)	(0.034)	(0.039)	(0.034)	(0.024)	
Presence x 2011	0.122^{***}	0.013^{*}		0.009	-0.069^{**}	0.098^{***}	0.097^{***}	-0.128^{***}	0.105^{***}
	(0.046)	(0.007)		(0.006)	(0.031)	(0.036)	(0.034)	(0.024)	(0.023)
Presence x 2016	0.151^{***}	0.019^{**}	0.063^{***}	0.011^{*}	-0.141^{***}	0.146^{***}	0.099^{***}	-0.091^{***}	0.049^{**}
	(0.053)	(0.008)	(0.016)	(0.006)	(0.032)	(0.037)	(0.036)	(0.026)	(0.020)
Presence x 2020	0.289^{***}	0.035^{***}	0.095^{***}	-0.001					
	(0.074)	(0.010)	(0.016)	(0.005)					
Diff 2016-2011	0.029	0.006		0.002	-0.072	0.048	0.001	0.037	-0.056
SE Diff 2016-2011	0.016	0.002		0.001	0.010	0.012	0.008	0.010	0.019
Diff 2020-2011	0.167	0.023		-0.010					
SE Diff 2020-2011	0.046	0.005		0.003					
Diff 2020-2016	0.138	0.016	0.032	-0.012	-0.012	-0.012	-0.012	-0.012	
SE Diff 2020-2016	0.036	0.003	0.014	0.003	0.003	0.003	0.003	0.003	
Presence Measure	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20
Sample Distance (km)	100	100	100	100	100	100	100	100	100
Controls x Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	14015	14015	8409	14015	11212	11212	11212	11212	13738
R^2 (full model)	0.854	0.844	0.737	0.870	0.538	0.642	0.776	0.864	0.196
R^2 (proj model)	0.062	0.067	0.117	0.202	0.114	0.049	0.135	0.066	0.054
Num. groups: parish_id	2803	2803	2803	2803	2803	2803	2803	2803	
Num. groups: region_year	20	20	12	20	16	16	16	16	
Num. groups: region									4
Num. groups: year									3

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table S9: Regression table for Public Goods outcomes: For parishes within 100km, OLS models interacting Nearest + 20km refugee presence (time-varying) with year, interacting controls with year, and including parish FE and region-year FE, with standard errors clustered at the parish level.

	Public Pri	Public Sec	Roads	PG Index	HC 2	HC 3	HC 4	HC 5	Health Util
Baseline Presence	-0.076	-0.010	0.040^{*}	-0.014^{***}	0.109^{***}	-0.138^{***}	-0.219^{***}	0.065^{***}	-0.166^{***}
	(0.057)	(0.006)	(0.022)	(0.005)	(0.035)	(0.033)	(0.033)	(0.022)	(0.013)
Presence x 2006	0.137^{***}	0.008^{*}		0.017^{***}	-0.193^{***}	0.116^{***}	0.204^{***}	-0.136^{***}	
	(0.051)	(0.005)		(0.005)	(0.033)	(0.033)	(0.030)	(0.020)	
Presence x 2011	0.147^{***}	0.009		0.024^{***}	-0.106^{***}	0.121^{***}	0.217^{***}	-0.099^{***}	0.142^{***}
	(0.053)	(0.006)		(0.005)	(0.029)	(0.029)	(0.029)	(0.020)	(0.017)
Presence x 2016	0.207^{***}	0.011	0.022	0.021^{***}	-0.195^{***}	0.167^{***}	0.192^{***}	-0.058^{***}	0.100^{***}
	(0.064)	(0.007)	(0.015)	(0.005)	(0.030)	(0.029)	(0.030)	(0.021)	(0.016)
Presence x 2020	0.327^{***}	0.028^{***}	0.047^{***}	0.006					
	(0.081)	(0.009)	(0.015)	(0.005)					
Diff 2016-2011	0.060	0.001		-0.004	-0.089	0.046	-0.024	0.041	-0.042
SE Diff 2016-2011	0.019	0.002		0.001	0.010	0.010	0.007	0.009	0.015
Diff 2020-2011	0.180	0.019		-0.018					
SE Diff 2020-2011	0.045	0.004		0.003					
Diff 2020-2016	0.120	0.018	0.025	-0.015	-0.015	-0.015	-0.015	-0.015	
SE Diff 2020-2016	0.032	0.003	0.013	0.003	0.003	0.003	0.003	0.003	
Presence Measure	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20
Sample Distance (km)	200	200	200	200	200	200	200	200	200
Controls x Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	22885	22885	13731	22885	18308	18308	18308	18308	24102
R^2 (full model)	0.872	0.848	0.732	0.865	0.603	0.680	0.770	0.874	0.174
R^2 (proj model)	0.034	0.046	0.085	0.160	0.120	0.054	0.102	0.057	0.058
Num. groups: parish_id	4577	4577	4577	4577	4577	4577	4577	4577	
Num. groups: region_year	20	20	12	20	16	16	16	16	
Num. groups: region									4
Num. groups: year									3

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table S10: Regression table for Public Goods outcomes: For parishes within 200km, OLS models interacting Nearest + 20km refugee presence (time-varying) with year, interacting controls with year, and including parish FE and region-year FE, with standard errors clustered at the parish level.

	Public Pri	Public Sec	Roads	PG Index	HC 2	HC 3	HC 4	HC 5	Health Util
Baseline Presence	-0.069	-0.009	0.032	-0.008^{*}	0.130^{***}	-0.122^{***}	-0.184^{***}	0.076^{***}	-0.094^{***}
	(0.053)	(0.006)	(0.022)	(0.005)	(0.034)	(0.032)	(0.032)	(0.022)	(0.012)
Presence x 2006	0.115^{***}	0.007		0.012^{**}	-0.214^{***}	0.117^{***}	0.176^{***}	-0.150^{***}	
	(0.042)	(0.004)		(0.005)	(0.033)	(0.032)	(0.029)	(0.020)	
Presence x 2011	0.129^{***}	0.008		0.018^{***}	-0.123^{***}	0.114^{***}	0.180^{***}	-0.118^{***}	0.100^{***}
	(0.046)	(0.006)		(0.005)	(0.028)	(0.029)	(0.029)	(0.020)	(0.016)
Presence x 2016	0.185^{***}	0.009	0.038^{***}	0.015^{***}	-0.212^{***}	0.169^{***}	0.159^{***}	-0.080^{***}	0.059^{***}
	(0.055)	(0.007)	(0.014)	(0.005)	(0.030)	(0.029)	(0.029)	(0.021)	(0.015)
Presence x 2020	0.304^{***}	0.027^{***}	0.059^{***}	0.002					
	(0.074)	(0.009)	(0.014)	(0.005)					
Diff 2016-2011	0.057	0.001		−0.003	−0.089	0.055	−0.021	0.038	−0.042
SE Diff 2016-2011	0.018	0.002		0.001	0.010	0.011	0.007	0.009	0.014
Diff 2020-2011	0.176	0.019		−0.016					
SE Diff 2020-2011	0.044	0.004		0.003					
Diff 2020-2016	0.119	0.018	0.021	−0.013	−0.013	−0.013	−0.013	−0.013	
SE Diff 2020-2016	0.032	0.003	0.012	0.003	0.003	0.003	0.003	0.003	
Presence Measure	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20
Sample Distance (km)	All	All	All	All	All	All	All	All	All
Controls x Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	25835	25835	15501	25835	20668	20668	20668	20668	30704
R^2 (full model)	0.870	0.847	0.725	0.871	0.598	0.718	0.779	0.879	0.157
R^2 (proj model)	0.029	0.042	0.077	0.161	0.118	0.051	0.095	0.056	0.050
Num. groups: parish_id	5167	5167	5167	5167	5167	5167	5167	5167	
Num. groups: region_year	20	20	12	20	16	16	16	16	
Num. groups: region									4
Num. groups: year									3

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table S11: Regression table for Public Goods outcomes: For all parishes, OLS models interacting Nearest + 20km refugee presence (time-varying) with year, interacting controls with year, and including parish FE and region-year FE, with standard errors clustered at the parish level.

	Neighbors	Move Freely	Felt Unsafe	Feared Crime	Born Non-Ugandan	Naturalize
Baseline Presence	0.046	-0.031	0.013	-0.021	-0.056	0.069^{*}
	(0.030)	(0.038)	(0.027)	(0.036)	(0.037)	(0.037)
Presence x 2006				-0.106^{*}		
				(0.054)		
Presence x 2011				0.048		
				(0.051)		
Presence x 2016	-0.011		-0.020	0.135^{**}		
	(0.067)		(0.061)	(0.067)		
Presence x 2020	-0.004	0.063	-0.073	-0.100		
	(0.062)	(0.066)	(0.057)	(0.063)		
Diff 2016-2011				0.087		
SE Diff 2016-2011				0.068		
Diff 2020-2011				−0.148		
SE Diff 2020-2011				0.063		
Diff 2020-2016	0.006		−0.052	−0.235		
SE Diff 2020-2016	0.080		0.072	0.077		
Presence Measure	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20
Sample Distance (km)	100	100	100	100	100	100
Controls x Year	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	2141	1556	2134	4500	1084	1066
R^2 (full model)	0.067	0.024	0.028	0.077	0.034	0.077
R^2 (proj model)	0.031	0.018	0.024	0.064	0.020	0.030
Num. groups: region	4	3	4	4	3	3
Num. groups: year	3	2	3	5		

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table S12: Regression table for Afrobarometer outcomes: For respondents within 100km, OLS models interacting Nearest + 20km refugee presence (time-varying) with year, interacting controls with year, and including region and year FE.

	Neighbors	Move Freely	Felt Unsafe	Feared Crime	Born Non-Ugandan	Naturalize
Baseline Presence	0.068^{***}	-0.036	0.047^{**}	-0.029	-0.083^{***}	0.071^{**}
	(0.025)	(0.031)	(0.024)	(0.026)	(0.028)	(0.028)
Presence x 2006				0.006		
				(0.041)		
Presence x 2011				0.034		
				(0.037)		
Presence x 2016	0.005		-0.129^{***}	0.256^{***}		
	(0.047)		(0.045)	(0.047)		
Presence x 2020	-0.007	0.084	-0.092^{**}	-0.074		
	(0.048)	(0.052)	(0.046)	(0.048)		
Diff 2016-2011				0.222		
SE Diff 2016-2011				0.047		
Diff 2020-2011				-0.108		
SE Diff 2020-2011				0.048		
Diff 2020-2016	-0.012		0.037	-0.331		
SE Diff 2020-2016	0.058		0.056	0.057		
Presence Measure	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20
Sample Distance (km)	200	200	200	200	200	200
Controls x Year	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	3977	2807	3971	8241	1925	1898
R^2 (full model)	0.049	0.016	0.021	0.040	0.031	0.077
R^2 (proj model)	0.024	0.012	0.018	0.036	0.024	0.034
Num. groups: region	4	4	4	4	4	4
Num. groups: year	3	2	3	5		

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table S13: Regression table for Afrobarometer outcomes: For respondents within 200km, OLS models interacting Nearest + 20km refugee presence (time-varying) with year, interacting controls with year, and including region and year FE.

	Neighbors	Move Freely	Felt Unsafe	Feared Crime	Born Non-Ugandan	Naturalize
Baseline Presence	0.034	-0.044	-0.026	-0.022	-0.076^{***}	0.073^{***}
	(0.023)	(0.030)	(0.023)	(0.026)	(0.027)	(0.027)
Presence x 2006				-0.014		
				(0.040)		
Presence x 2011				-0.069^{*}		
				(0.036)		
Presence x 2016	0.046		-0.054	0.268^{***}		
	(0.044)		(0.044)	(0.046)		
Presence x 2020	0.056	0.085^{*}	-0.043	-0.091^{**}		
	(0.044)	(0.048)	(0.044)	(0.046)		
Diff 2016-2011				0.337		
SE Diff 2016-2011				0.045		
Diff 2020-2011				-0.022		
SE Diff 2020-2011				0.044		
Diff 2020-2016	0.010		0.011	-0.359		
SE Diff 2020-2016	0.054		0.055	0.054		
Presence Measure	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20
Sample Distance (km)	All	All	All	All	All	All
Controls x Year	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	4699	3341	4695	9499	2334	2305
R^2 (full model)	0.045	0.013	0.019	0.042	0.026	0.075
R^2 (proj model)	0.019	0.010	0.015	0.042	0.019	0.028
Num. groups: region	4	4	4	4	4	4
Num. groups: year	3	2	3	5		

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table S14: Regression table for Afrobarometer outcomes: For all respondents, OLS models interacting Nearest + 20km refugee presence (time-varying) with year, interacting controls with year, and including region and year FE.

S2.3 Including a District Time Trend

These models include an additional district time trend.

	Public Pri	Public Sec	Roads	PG Index	HC 2	HC 3	HC 4	HC 5	Health Util
Baseline Presence	0.044	-0.003	0.053^{**}	-0.038^{***}	-0.050	-0.205^{***}	-0.273^{***}	0.112^{***}	-0.107^{***}
	(0.084)	(0.007)	(0.024)	(0.006)	(0.036)	(0.035)	(0.034)	(0.023)	(0.014)
Presence x 2006	0.122^{**}	0.005		0.022^{***}	-0.149^{***}	0.115^{***}	0.187^{***}	-0.159^{***}	
	(0.061)	(0.004)		(0.005)	(0.032)	(0.034)	(0.032)	(0.021)	
Presence x 2011	0.052	0.005		0.036^{***}	0.029	0.144^{***}	0.190^{***}	-0.100^{***}	0.107^{***}
	(0.061)	(0.006)		(0.005)	(0.029)	(0.033)	(0.034)	(0.022)	(0.020)
Presence x 2016	0.079	0.011	0.038^{**}	0.032^{***}	0.001	0.182^{***}	0.137^{***}	-0.025	0.065^{***}
	(0.070)	(0.008)	(0.017)	(0.006)	(0.034)	(0.039)	(0.042)	(0.029)	(0.022)
Presence x 2020	0.161^{**}	0.024^{***}	0.039	0.014^{**}					
	(0.076)	(0.009)	(0.024)	(0.006)					
Presence Measure	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20	Nearest+20
Sample Distance (km)	150	150	150	150	150	150	150	150	150
Controls x Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	18185	18185	10911	18185	14548	14548	14548	14548	17982
R^2 (full model)	0.875	0.854	0.744	0.874	0.598	0.687	0.801	0.879	0.201
R^2 (proj model)	0.039	0.044	0.048	0.172	0.083	0.041	0.071	0.053	0.068
Num. groups: parish_id	3637	3637	3637	3637	3637	3637	3637	3637	
Num. groups: region_year	20	20	12	20	16	16	16	16	
Num. groups: year_n:as.factor(district)	45	45	45	45	45	45	45	45	41
Num. groups: region									4

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table S15: Regression table for Public Goods outcomes: For respondents within 150km, OLS models interacting Nearest + 20km refugee presence (time-varying) with year, interacting controls with year, and including parish FE, region-year FE, and a district time trend, with standard errors clustered at the parish level.

S2.4 Two Period Binary Treatment DiD

This section shows the results for public goods outcomes using a standard DiD specification: two period (pre-2014 and post-2014) with a binary measure of refugee presence: Nearest + 20km cutoff at the median.

	Public Pri	Public Sec	Roads	HC2	HC3	HC4	HC5	PG Index
Exposure	0.058^{**}	-0.012	0.179^{***}	-0.160^{***}	0.110^{***}	-0.061^{**}	0.068^{**}	0.031^{***}
	(0.025)	(0.029)	(0.032)	(0.034)	(0.027)	(0.025)	(0.028)	(0.005)
Controls x Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Distance (km)	150	150	150	150	150	150	150	150
Num. obs.	18178	18178	10904	14548	14548	14548	14548	18178
R^2 (full model)	0.944	0.873	0.716	0.535	0.651	0.740	0.851	0.857
R^2 (proj model)	0.042	0.050	0.127	0.124	0.038	0.063	0.066	0.629
Num. groups: time_fe	2	2	2	2	2	2	2	2
Num. groups: parish_id	3637	3637	3637	3637	3637	3637	3637	3637
Num. groups: region	4	4	4	4	4	4	4	4

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table S16: Regression table for Public Goods outcomes: For parishes within 150km, OLS models using Nearest + 20km binary refugee presence in a two-period model (pre-2014 vs. post-2014), including controls, parish FE and region FE, with standard errors clustered at the parish level.

S3 Sensitivity Analysis

In this section, we implement sensitivity analysis from Cinelli and Hazlett (2020) to understand the impact of omitted variables.

outcome	estimate	se	t_statistic	r2yd.x	rv_q	rv_qa	adj_est_1x	adj_est_5x	adj_est_10x
Public Primary	0.34	0.08	4.10	0.5%	6.6%	3.5%	0.33	0.30	0.27
Public Secondary	0.04	0.01	4.14	0.5%	6.6%	3.6%	0.03	0.01	-0.01
Road Density	0.08	0.01	5.50	0.8%	8.7%	5.7%	0.08	0.08	0.08
PG Index	0.02	0.00	3.18	0.3%	5.1%	2.0%	0.01	0.01	-0.00
HC2	-0.14	0.03	-4.71	0.6%	7.5%	4.5%	-0.12	-0.06	0.02
HC3	0.17	0.03	5.58	0.8%	8.8%	5.8%	0.17	0.14	0.12
HC4	0.23	0.03	7.32	1.5%	11.4%	8.5%	0.23	0.20	0.17
HC5	-0.06	0.02	-2.80	0.2%	4.5%	1.4%	-0.06	-0.02	0.03

This table shows the sensitivity diagnostics for the DiD estimates for refugee presence interacted with 2020. To interpret these sensitivity analyses:

- Partial R2 of the treatment with the outcome (`r2yd.x`): an extreme confounder (orthogonal to the covariates) that explains 100% of the residual variance of the outcome, would need to explain at least X% of the residual variance of the treatment to fully account for the observed estimated effect.

- Robustness Value, q = 1 (`rv_q`): unobserved confounders (orthogonal to the covariates) that explain more than X% of the residual variance of both the treatment and the outcome are strong enough to bring the point estimate to 0 (a bias of 100% of the original estimate). Conversely, unobserved confounders that do not explain more than X% of the residual variance of both the treatment and the outcome are not strong enough to bring the point estimate to 0.

- Robustness Value, q = 1, alpha = 0.05 (`rv_qa`): unobserved confounders (orthogonal to the covariates) that explain more than X% of the residual variance of both the treatment and the outcome are strong enough to bring the estimate to a range where it is no longer 'statistically different' from 0 (a bias of 100% of the original estimate), at the significance level of alpha = 0.05. Conversely, unobserved confounders that do not explain more than X% of the residual variance of both the treatment and the outcome are not strong enough to bring the estimate to a range where it is no longer 'statistically different' from 0, at the significance level of alpha = 0.05.

- `adj_est_1x` / `adj_est_5x` / `adj_est_10x`: what the adjusted estimate in the presence of an unobserved confounder that is as large as / 5 times as large as / 10 times as large as the interaction between parish unemployment rate measured in the 2002 census interacted with 2016.

Our results are robust to concerns of omitted variables bias.

S4 Multiple Hypothesis Testing Corrections

We address concerns about multiple hypothesis testing by adjusting for the false discovery rate (FDR). We show the Benjamini-Hochberg (BH) adjusted p-values for the results in our paper. We show the adjusted p-values within sets of outcomes for the 2020 estimates. Our statistically significant findings for the main public goods outcomes hold.

S4.1 Public Goods Outcomes

Outcome	Estimate	SE	p-value	adj. p-value
Public Primary	0.335	0.082	0.000	0.000
Public Secondary	0.036	0.009	0.000	0.000
Road Density	0.081	0.015	0.000	0.000
PG Index	0.016	0.005	0.001	0.001

S4.2 Migration Attitudes (Afrobarometer)

Outcome	Estimate	SE	p-value	adj. p-value
Migrant Neighbors	0.071	0.027	0.008	0.015
Migrant Neighbors	-0.006	0.053	0.911	0.911
Move Freely	-0.019	0.032	0.563	0.676
Move Freely	0.041	0.055	0.453	0.676
Born non-Ugandan	-0.099	0.032	0.002	0.011
Naturalize	0.092	0.032	0.004	0.011

S4.3 Perceptions of Insecurity (Afrobarometer)

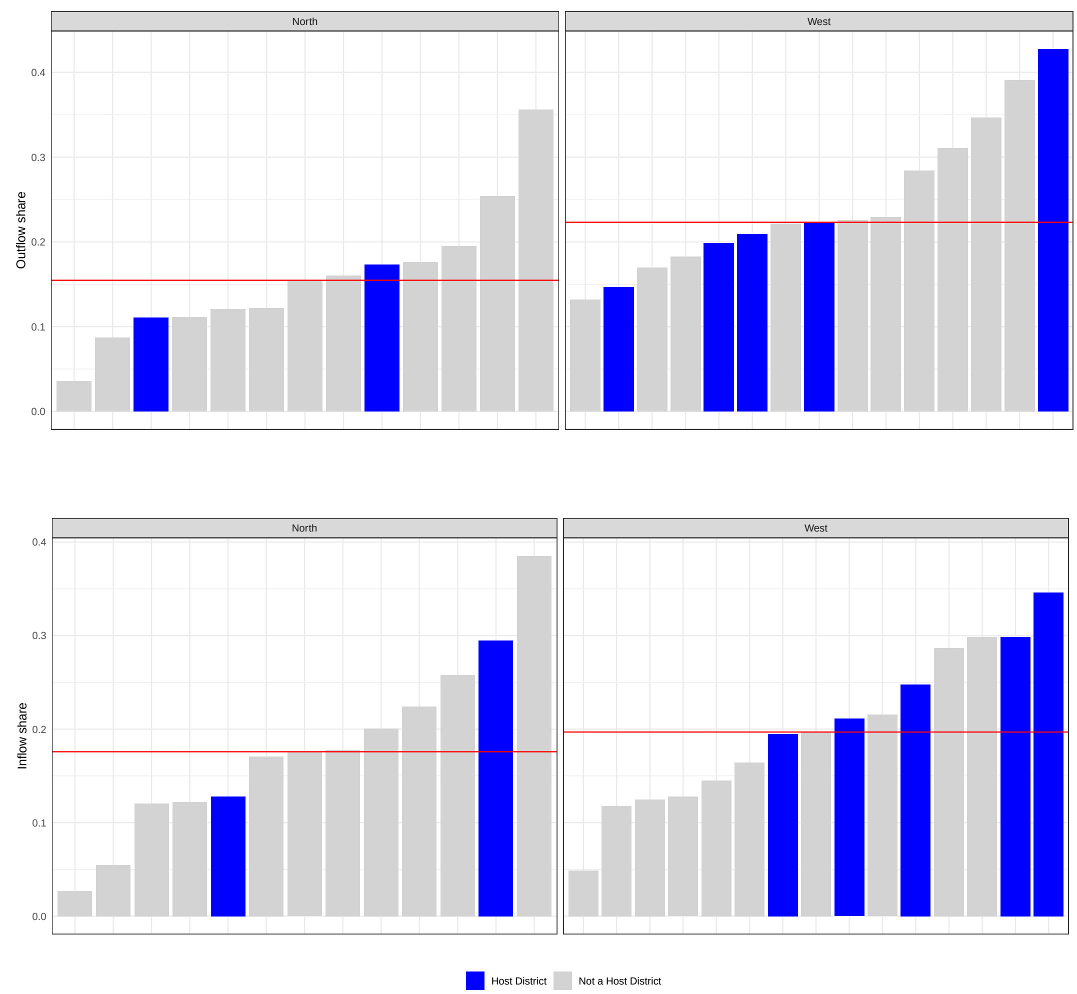
Outcome	Estimate	SE	p-value	adj. p-value
Feel Unsafe	-0.109	0.049	0.026	0.026
Fear Crime	-0.130	0.053	0.013	0.026

S5 Addressing Compositional Concerns of Domestic Migration

In order to rule out the possibility that our findings on public opinion are not driven potentially by population re-composition, we provide additional evidence by analyzing Ugandans' mobility on district level. We used geocoded individual level Uganda DHS (2016) data that specifically asked questions about respondent's previous and current district of residence, based on which we generated a dummy variable that takes 1 if the respondent moved in current residential district from a different one and 0 otherwise.

Firstly, we aggregated individual mobility indicator to district level and calculated an inflow and outflow ratio for each district. Outflow ratio is defined as the number of people that moved out of their previous district divided by the original total population of that district. Inflow ratio is defined in a similar way. Figure S13 displays the outflow and inflow ratios for districts in the North and West region where refugee camps are located. Compared to those that do not host refugees, refugee-hosting districts have relatively low out-migration mobility while have in-migration mobility close and some higher than the median.

Figure S13: Outflow & Inflow ratio for districts in the North and West region.



While we cannot assess comprehensive reasons why people migrated into these refugee hosting districts due to lack of supportive questions in DHS questionnaires, we applied simple linear regression model with controls and region fixed effects on individual level to examine whether refugee presence contributed to in-migration mobility. One very important note is that there is high probability that Uganda DHS 2016 sample refugees together with Ugandans and didn't provide information that could help distinguish the two (compare the sampling method description for 2016 DHS versus 2018 DHS where they specifically emphasized and distinguished refugee respondents (UBOS, 2018; Uganda National Malaria Control Division , NMCD)). In addition, an unusually large wave of internal displacement due to conflict and violence was reported happening in 2016 (Centre, 2021), as well as refugee displacement (Agency, 2016). Taking these factors into account, we use a radius cutoff based on distance to refugee camps In order to restrain the sample to Ugandans only. More specifically, we have four different sets of cohorts for our regression analysis: 10 - 100km, 10 - 150km, 10 - 200km, all that are farther than 10km. In the following regression table, we show that there is no relationship between in-migration mobility and our refugee presence measure, whether in continuous or dichotomous scale.

	continuous presence measure				dichotomous presence measure			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nearest + 20km	0.010	0.009	0.012	0.007	0.003	0.028	−0.010	−0.004
	(0.0148)	(0.0123)	(0.0139)	(0.0142)	(0.0316)	(0.0249)	(0.0330)	(0.0323)
Sample Distance (km)	100	150	200	All	100	150	200	All
Num. obs.	9442	12178	17229	22083	9442	12178	17229	22083
N Clusters	31	39	47	55	31	39	47	55
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.030	0.056	0.081	0.089	0.029	0.056	0.081	0.089

Again, our analysis was carried out using 2002 administrative districts for the purpose of consistency. But note the we repeated the same analysis using 2016 districts. The results remain the same.

References

Adhikari, Ramesh. 2016. "Effect of Women's autonomy on maternal health service utilization in Nepal: a cross sectional study." *BMC women's health* 16(1):1–7.

Agency, UNHCR: The UN Refugee. 2016. "Global Trends: Forced Displacement in 2016.".
URL: *https://www.unhcr.org/5943e8a34.pdf*

Bowles, Jeremy, Horacio Larreguy and Anders Woller. 2020. "Information Versus Control: The Electoral Consequences of Polling Place Creation." *IAST working paper* .

Center For International Earth Science Information Network-CIESIN-Columbia University and Information Technology Outreach Services-ITOS-University Of Georgia. 2013. "Global Roads Open Access Data Set, Version 1 (gROADSv1).".
URL: *https://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1*

Centre, Internal Displacement Monitoring. 2021. "Uganda.".
URL: *https://www.internal-displacement.org/countries/uganda*

Cinelli, Carlos and Chad Hazlett. 2020. "Making sense of sensitivity: Extending omitted variable bias." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82(1):39–67.

Dimbuene, Zacharie Tsala, Joshua Amo-Adjei, Dickson Amugsi, Joyce Mumah, Chimaraoke O Izugbara and Donatien Beguy. 2018. "Women's education and utilization of maternal health services in Africa: a multi-country and socioeconomic status analysis." *Journal of biosocial science* 50(6):725–748.

Grossman, Guy and Janet I. Lewis. 2014. "Administrative Unit Proliferation." *The American Political Science Review* 108:196–217.

Gyimah, Stephen Obeng, Baffour K Takyi and Isaac Addai. 2006. "Challenges to the reproductive-health needs of African women: on religion and maternal health utilization in Ghana." *Social science & medicine* 62(12):2930–2944.

Magadi, Monica Akinyi, Alfred O Agwanda and Francis O Obare. 2007. "A comparative analysis of the use of maternal health services between teenagers and older mothers in sub-Saharan Africa: evidence from Demographic and Health Surveys (DHS)." *Social science & medicine* 64(6):1311–1325.

MOH. 2016. Services standards and service delivery standards for the health sector. Technical report Ministry of Health, Uganda.

Mpango, Jonathan and Josephine Nabukenya. 2019. A Qualitative Study to Examine Approaches used to Manage Data about Health Facilities and their Challenges: A Case of Uganda. In *AMIA Annual Symposium Proceedings*. Vol. 2019 American Medical Informatics Association p. 1157.

Müller-Crepon, Carl, Philipp Hunziker and Lars-Erik Cederman. 2021. "Roads to Rule, Roads to Rebel: Relational State Capacity and Conflict in Africa." *Journal of Conflict Resolution* 65(2-3):563–590.

UBOS. 2016. The National population and Housing Census 2014-Main Report. Technical report Uganda Bureau of Statistics Kampala, Uganda: .

UBOS, ICF. 2018. Uganda Demographic and Health Survey 2016. Technical report Uganda Bureau of Statistics and ICF Kampala, Uganda and Rockville, maryland, USA: .

Uganda National Malaria Control Division (NMCD), (UBOS) and ICF. 2020. Uganda Malaria Indicator Survey 2018-19. Technical report Uganda National Malaria Control Division (NMCD), Uganda Bureau of Statistics (UBOS), and ICF Kampala, Uganda and Rockville, maryland, USA: .

Yaya, Sanni et al. 2020. "Wealth status, health insurance, and maternal health care utilization in Africa: evidence from Gabon." *BioMed Research International* 2020.